Structure, Vol. 12, 1–20, September, 2004, ©2004 Elsevier Ltd. All rights reserved. DOI 10.1016/j.str.2004.06.016

Crystal Structure of *A. fulgidus* Rio2 Defines a New Family of Serine Protein Kinases

Nicole LaRonde-LeBlanc and Alexander Wlodawer
Protein Structure Section
Macromolecular Crystallography Laboratory
National Cancer Institute
NCI-Frederick
Frederick, Maryland 21702

Summary

The RIO family of atypical serine/threonine kinases contains two subfamilies, Rio1 and Rio2, highly conserved from archaea to man. Both RIO proteins from *Saccharomyces cerevisiae* catalyze serine phosphorylation in vitro, and the presence of conserved catalytic residues is required for cell viability. The activity of Rio2 is necessary for rRNA cleavage in 40S ribosomal subunit maturation. We solved the X-ray crystal structure of *Archaeoglobus fulgidus* Rio2, with and without bound nucleotides, at 2.0 Å resolution. The C-terminal RIO domain is indeed structurally homologous to protein kinases, although it differs from known serine kinases in ATP binding and lacks the regions important for substrate binding. Unexpectedly, the N-terminal Rio2-specific domain contains a winged helix fold, seen primarily in DNA-binding proteins. These discoveries have implications in determining the target and function of RIO proteins and define a distinct new family of protein kinases.

Introduction

Protein kinases are essential for the regulation of a large variety of cellular processes including cell cycle progression, transcription, DNA replication, and metabolic functions. The protein kinase superfamily represents one of the largest protein superfamilies in eukaryotes (Hanks and Hunter, 1995), with over 500 representatives in the human genome (Manning et al., 2002). Protein kinases can be divided into two major classes that catalyze phosphorylation of either serine and threonine, or tyrosine residues (Bossemeyer, 1995, 2002; Hanks et al., 1988). They are characterized by conserved catalytic domains, 250–300 amino acids in length, composed of conserved secondary structure elements and key sequences referred to as "subdomains" (Bossemeyer, 1995; Hanks et al., 1988; Hanks and Hunter, 1995). A number of structures of protein-serine/threonine and protein-tyrosine kinases have been determined, and the conserved subdomain residues have been shown to be involved in recognition and binding of ATP or substrate peptides, as well as in actual catalysis (Bossemeyer, 1995; Engh and Bossemeyer, 2002; Knighton et al., 1991a, 1991b).

Rio1 from *Saccharomyces cerevisiae* is the founding member of a recently described family of serine kinases that are highly conserved yet very divergent from the kinase families with known structures (Angermayr et al., 2002; Angermayr and Bandlow, 2002). RIO kinases have been classified as atypical protein kinases (aPKs) that are part of the human "kinome" (Manning et al., 2002). Through analysis of multiple protein sequence alignments and secondary structure predictions, RIO proteins were found to contain the conserved signature residues of protein kinases (Leonard et al., 1998; Angermayr and Bandlow, 2002). Such alignments also revealed that this family can be further subdivided into at least two subfamilies, Rio1 and Rio2 (Leonard et al., 1998; Vanrobays et al., 2003; Geerlings et al., 2003). While the Rio1 proteins contain the conserved RIO kinase domain, Rio2 proteins also contain an additional N-terminal domain of unknown function (Vanrobays et al., 2003). Representatives of each family are present in organisms from archaea to man, suggesting a fundamental role in the cell (Geerlings et al., 2003; Vanrobays et al., 2003). Both Rio1 and Rio2 have been shown to have serine kinase activity in vitro, and residues that are important in kinase domains for catalysis are necessary for cell viability (Geerlings et al., 2003; Angermayr and Bandlow, 2002; Vanrobays et al., 2001, 2003).

Yeast Rio1 plays a critical role in cell cycle progression (Angermayr et al., 2002). Studies have indicated that cells with impaired Rio1 function exhibit growth arrest in M phase or late G1 phase, implying that Rio1 is involved in regulation of entry into S phase and exit from mitosis. Both RIO proteins are critically important for ribosome biogenesis. Processing of 20S pre-rRNA to the 18S rRNA of the ribosomal 40S subunit absolutely requires both Rio1 and Rio2 in yeast (Schafer et al., 2003; Vanrobays et al., 2001, 2003). Although there is significant sequence similarity between Rio1 and Rio2 proteins (43% similarity between the yeast enzymes), Rio2 proteins are functionally distinct from Rio1 proteins and do not complement their activity, since deletion of Rio2 in yeast is lethal, despite functional Rio1 (Giaever et al., 2002). In yeast cells in which the enzymatic activity of Rio2 has been abrogated by mutation, 20S pre-rRNA accumulates both in the nucleus and in the cytoplasm, suggesting that enzymatic activity is necessary for cleavage but not for export of the rRNA (Geerlings et al., 2003). Rio2 has been found in yeast protein complexes, isolated by tandem-affinity purification with many proteins involved in ribosome biogenesis, as well as many proteins involved in cell proliferation (Schafer et al., 2003; Gavin et al., 2002; Ho et al., 2002). It is not currently known how the roles of RIO proteins in cell cycle progression and ribosome maturation are connected, although it has been shown that interruption of ribosome biogenesis can result in cell cycle arrest (Ruggero and Pandolfi, 2003).

In order to gain insight into the function of these very divergent kinases, we determined the X-ray crystal structure of the full-length Rio2 protein kinase from *Archaeoglobus fulgidus*. The structure shows a kinase domain that is truncated and more compact than canon-

*Correspondence: wlodawer@ncifcrf.gov

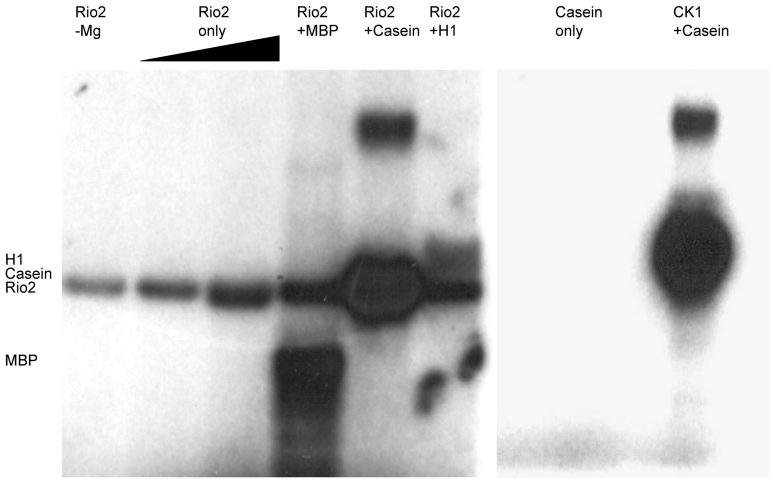

Figure 1. AfRio2 Is an Active Kinase

Autoradiogram of ^{32}P-labeled Rio2, myelin basic protein, casein, and histone H1. Autophosphorylation in the absence (Rio2 − Mg) and presence of magnesium (Rio2 only); phosphorylation of common kinase substrates myelin basic protein, α-casein, and histone H1 (Rio2 + MBP, Rio2 + casein, Rio2 + H1, respectively); control lanes of α-casein in the absence and presence of added casein kinase (casein only and CK1 + casein).

ical protein kinases. The RIO kinase domain contains all of the residues absolutely required for catalysis, but none of the subdomains known to be necessary in canonical protein kinases for substrate binding outside of the active site. Surprisingly, the Rio2-specific N-terminal domain is a winged helix domain, previously found mostly in DNA binding proteins and directly involved in nucleic acid binding. Therefore, the structure of Rio2 defines a structurally unique group of serine kinases containing a putative DNA binding domain and provides a framework for studying the function of this ancient protein family fundamental to archaeal and eukaryotic life.

Results

Structure Determination and the Overall Fold

Full-length Rio2 from *A. fulgidus* (AfRio2) was expressed in *E. coli* and purified using heat denaturation, anion exchange chromatography, and affinity chromatography. Mass spectrometry confirmed that the purified protein contained all the residues (1–282) and was unmodified (data not shown). Assays using the purified protein confirmed that the preparation was active, as shown by the ability to autophosphorylate as well as to phosphorylate supplied substrates in the presence of magnesium ions (Figure 1). Surprisingly, some autophosphorylation was also detected in the absence of magnesium. The protein was determined by analytical ultracentrifugation to be monomeric in solution in the absence of nucleotides (data not shown).

We obtained crystals of AfRio2, which diffract to ~2.0 Å and belong to the space group C2, with one molecule per asymmetric unit. The structure was solved using the multiple-wavelength anomalous diffraction (MAD) phasing technique with seleno-methionine (SeMet)-substituted protein at 1.99 Å. The model contains 263 of the 282 residues of AfRio2. Seventeen residues (between residues 124 and 142) are largely unstructured, although four of them were found packed against a neighboring symmetry-related molecule. The structure of the apoenzyme was refined to an R factor of 16.5% and an R_{free} of 19.2%. Data collection and crystallographic refinement statistics are shown in Table 1.

Analysis of the structure reveals the presence of two domains. The N-terminal domain, conserved in members of the Rio2 family and not present in the Rio1 family, is structurally homologous to the winged helix (wHTH) domain (Figure 2). In AfRio2, this domain contains four α helices followed by two β strands and a fifth α helix. A strand that connects the second and third α helices combines with the other two β strands to form a β sheet, and the loop between the second and third β strands is called a "wing," which gives the wHTH its name. The fifth α helix replaces the second wing seen in some proteins with the wHTH fold. The most commonly reported function of such domains is DNA binding (Gajiwala et al., 2000; Gajiwala and Burley, 2000). The C-terminal RIO domain, the sequence of which is conserved in both Rio1 and Rio2 proteins, is indeed structurally homologous to known protein kinase domains. It is bilobal with a twisted five-stranded β sheet and a single α helix in the N-terminal lobe (Figures 2 and 3), whereas the C-terminal lobe consists of a combination of four β strands and three α helices (Figures 2 and 4). No analogous combination of a wHTH domain and a kinase domain has been previously reported.

Comparison of the wHTH Domain with Its Structural Equivalents

The nearest structural neighbors to the AfRio2 wHTH domain, found with the DALI server (Holm and Sander, 1993), include transcription factors MarR and SlyA and histone linker protein GH5. A structure-based sequence alignment of the AfRio2 N-terminal domain and the wHTH domains of these proteins is shown in Figure 3C. Although these domains are structurally similar (rms deviations in the range of 1.2–1.3 Å), they share only 7% to 14% identity. Such domains are usually found in DNA binding proteins and have been shown to bind DNA in two different ways (Gajiwala and Burley, 2000; Wolberger and Campbell, 2000). The most common DNA binding mode, first reported in the crystal structure of transcription factor HNF-3 bound to DNA, is through site-specific interactions in the major grooves with resi-

Table 1. Data Collection and Refinement Statistics for the Apo, AMPPNP-, and ATP-Bound Rio2

Crystal Data: Space Group C2						
		Se-Met MAD		Apo	AMPPNP	ATP
a (Å)		116.65		116.59	116.17	116.14
b (Å)		44.62		44.45	44.23	44.51
c (Å)		62.58		62.60	62.10	62.68
β (°)		94.18		94.31	94.38	93.95
	Peak	Edge	Remote			
λ (Å)	0.97931	0.97942	0.99998	0.99998	0.99998	1.54
Resolution (Å)	20–2.00	20–2.00	20–2.00	50–2.00	29–1.99	50–2.10
R_{sym}(last shell)	0.035 (0.114)	0.035 (0.125)	0.039 (0.156)	0.049 (0.117)	0.042 (0.125)	0.052 (0.174)
Reflections	42,028 (4181)	42,140 (4169)	42,058 (3757)	21,731 (2129)	21,420 (2079)	18,666 (1831)
Redundancy	3.9 (3.8)	3.9 (3.7)	3.8 (3.2)	3.8 (3.7)	7.5 (6.8)	4.1 (4.1)
Completeness (%)	98.8 (98.0)	98.8 (97.9)	97.6 (86.4)	99.0 (98.1)	98.3 (95.7)	98.6 (97.8)
R/R_{free} (%)				16.7/21.2	17.1/22.8	16.5/21.5
(Last shell)				(15.7/26.0)	(18.2/28.5)	(18.3/28.5)
Mean B factor (Å^2)				23.14	23.56	31.01
Waters				205	198	172
Rms deviations						
Lengths (Å)				0.023	0.028	0.032
Angles (°)				2.07	2.06	2.45

dues on α helix 3 (Clark et al., 1993). The second mode, characterized by the crystal structure of hRFX1 bound to DNA, involves specific interactions between W1, β2, and β3 with the major groove of the DNA (Gajiwala et al., 2000). Among the top ten hits on the DALI server for the Rio2 domain, MarR and SlyA are likely to bind DNA via helix 3 in the major groove like HNF-3, while histone H5 is likely to bind in a manner analogous to hRFX1 (Alekshun et al., 2001; Wolberger and Campbell, 2000). Comparison of the structure of Rio2 with histone H5 and SlyA showed that Rio2 has more in common with latter than with the former (Figure 3). Like SlyA (and MarR; data not shown), Rio2 has a fourth helix rather than a second wing (Figure 3A). Analysis of the electrostatic surfaces of Rio2, histone H5, and SlyA shows the electrostatic properties of the surface that would bind DNA are most similar between SlyA and Rio2 (Figure 3B). Sequence analysis of Rio2 shows the presence of charged amino acids in helix 3 of the wHTH domain, which may play a role in DNA binding (Figure 3C). Initial experiments have indicated that Rio2 may bind nonspecifically to single-stranded oligonucleotides (unpublished data). Taken together, these data suggest Rio2 could potentially bind an oligonucleotide.

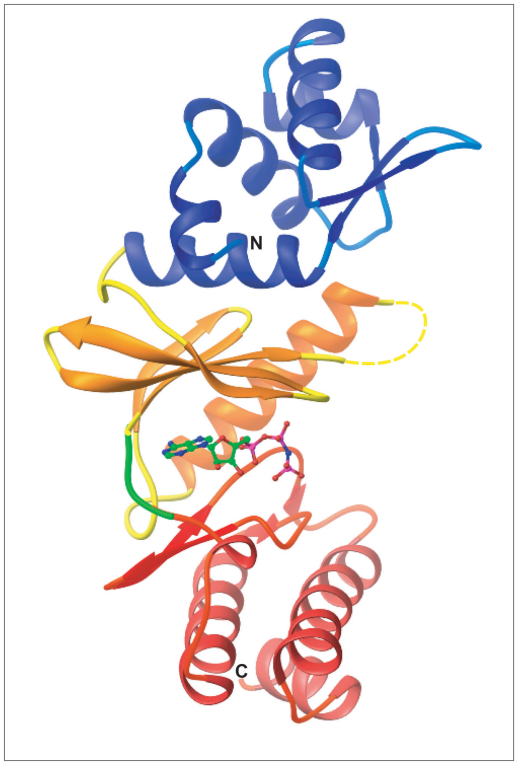

Figure 2. Overall Structure of Rio2

The crystal structure of AfRio2 shows a bilobal kinase domain (orange and red) tightly connected to an N-terminal winged helix domain (blue). The green "hinge" region connects the N- and C-terminal lobes of the kinase domain. The bound ATP analog, AMPPNP, is shown in a ball-and-stick representation. This figure was made using Ribbons (Carson, 1991).

Unique Characteristics of the RIO Kinase Domain

The structure of the Rio2 kinase domain was extensively compared with the three-dimensional structures of the catalytic domains of several protein kinases. The closest structural homolog to Rio2 identified by the DALI server was casein kinase (CK1) (rms deviation 2.8 Å). As shown in Figure 4, the RIO kinase domain retains most of the structural characteristics of protein kinases in alignments with CK1 and cyclic-AMP-dependent kinase (PKA). The N-terminal lobe is the typical five-stranded β sheet and long α helix C. A loop of 17 amino acids is disordered between β3 and αC (Figures 2 and 4). Disorder in this region was also observed in the structure of checkpoint kinase Chk1 and cyclin-dependent kinase 2 (CDK2) (Chen et al., 2000; Brown et al., 1999). In CDK2, this region is only ordered when the protein is bound by cyclin A, a regulating cofactor required for its activity. It should be noted that in Rio2 this region contains a highly conserved threonine and an adjacent serine (Figure 4B), possible targets for phosphorylation.

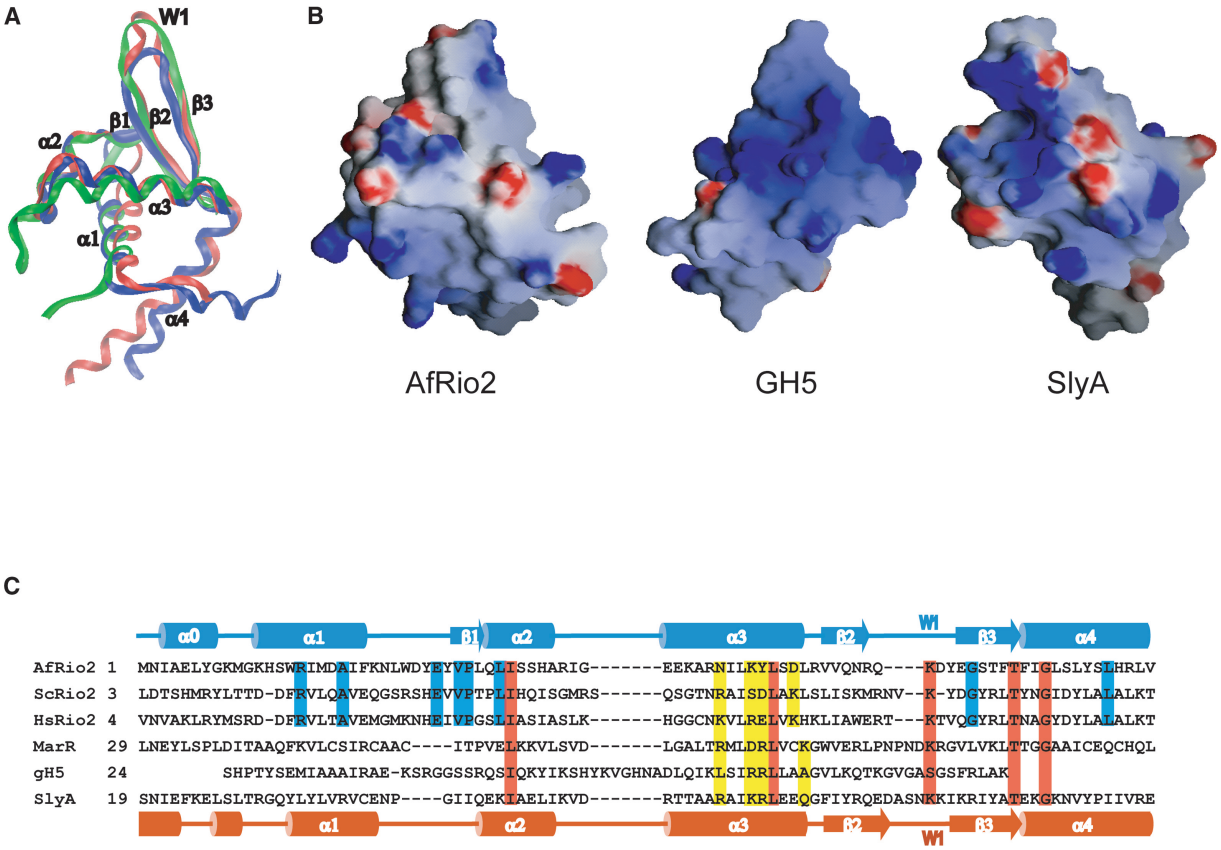

Figure 3. The Rio2 Winged Helix Domain

(A) Alignment of the wHTH domain of Rio2 (blue) with that of histone H5 (green) and SlyA (red) shows close similarity.

(B) Electrostatic surfaces of the three proteins in the same orientation as in (A) (−10 to +10 κBt, red to blue).

(C) Structure-based sequence alignment of the N-terminal domain of Rio2 from *A. fulgidus* (AfRio2), *S. cerevisiae* (ScRio2), and *H. sapiens* (HsRio2) with MarR, histone H5 (gH5), and SlyA. Residues identical in all the sequences except gH5 are highlighted in red, while blue highlighted residues are identical in Rio2 proteins only. Yellow highlights indicate residues that may be involved in DNA binding in the Rio2 proteins, MarR, and SlyA. The figures were made using VMD (Humphrey et al., 1996) and Grasp (Nicholls, 1992).

The several hundred known protein kinases contain 11 conserved subdomains that make up the catalytic core. Surprisingly, upon examination of the structure of the RIO kinase domain, only eight of these subdomains were found (Figure 4). The RIO kinase domain seen here is shorter than the canonical kinase domains, containing only two of the C-terminal lobe's usual five α helices (subdomains VIa and IX) and lacking the activation loop (subdomain VIII). This is clearly shown in structural alignments of Rio2 with PKA and CK1 (Figure 4). In crystal structures of several protein kinases complexed with peptide inhibitors, the activation loop and the helices of subdomains X and XI have been shown to create a surface for substrate binding (Cox et al., 1994; Engh and Bossemeyer, 2002; Knighton et al., 1991b; Zheng et al., 1993).

The structures of AfRio2 bound to adenosine triphosphate (ATP) and β-imino-adenosine-5′-triphosphate (AMPPNP) were solved at 2.10 and 1.99 Å, respectively, using crystals soaked in a cryoprotectant containing the nucleotide and magnesium ions. Nucleotide binding in Rio2 is similar to that seen in PKA and CK1 with respect to the adenine ring (Knighton et al., 1991a; Xu et al., 1995). In all the structures, a conserved Glu (Glu121 in PKA) forms a hydrogen bond with the exocyclic amine of the adenine ring which packs into a hydrophobic pocket. In Rio2, as in PKA, the ribose ring is firmly held in place by two hydrogen bonds between the 2′- and 3′-hydroxyl groups and the backbone of a nearby residue (Tyr222 in Rio2, Glu170 in PKA) and a carboxylate from an invariant Glu (Glu186 in Rio2, Glu127 in PKA). In PKA, an additional hydrogen bond is provided to the 3′-hydroxyl by an Arg side chain from the inhibitory peptide (Knighton et al., 1991b). An invariant Lys (120 in Rio2, 72 in PKA) makes a hydrogen bond to the α-phosphate in all kinases (Bossemeyer, 1995; Engh and Bossemeyer, 2002; Hanks and Hunter, 1995). This lysine is assisted by the invariant glutamic acid residue of helix C, which interacts with the lysine to hold it in place. The β-phosphate is contacted in Rio2 by Ser104, which replaces the third glycine in the canonical Gly-X-Gly-X-X-Gly sequence (Gly55 in PKA). In PKA and CK1, the β-phosphate is contacted instead by the invariant lysine (Knighton et al., 1991a; Xu et al., 1995). The positioning of the γ-phosphate is important for catalysis, and in the AMPPNP-containing structure, the γ-phosphate is found in a single conformation, forming water-mediated hydrogen bonds with Asn223 (171 in PKA),

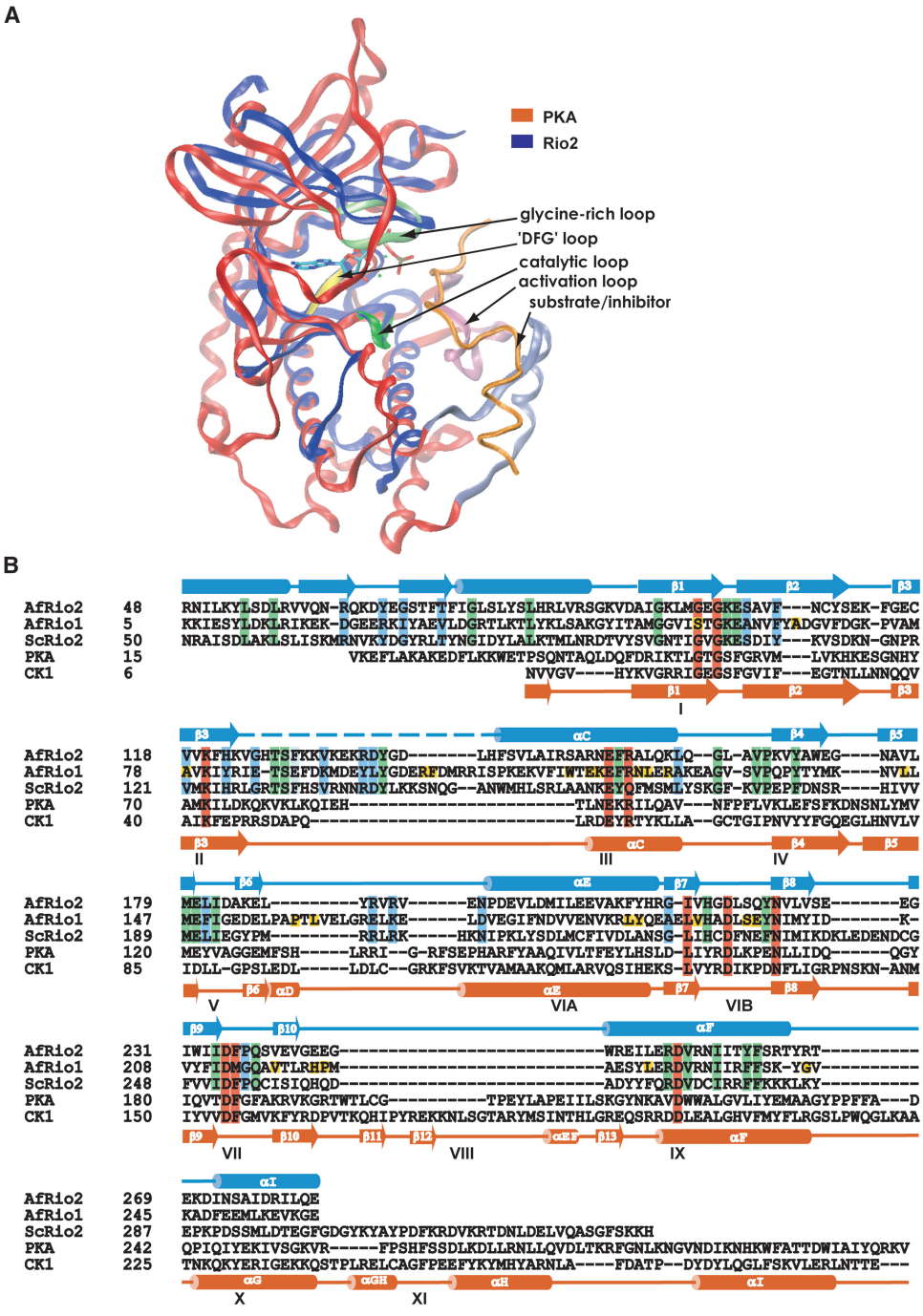

Figure 4. The RIO Kinase Domain

(A) Alignment of the nucleotide-bound structure of the Rio2 kinase domain (blue) and PKA (red with colored segments). The structures were aligned using the C-terminal lobe of each protein. The nucleotide shown is that bound to PKA. The mechanistically important parts of the PKA kinase domain are colored and labeled. The figure was made using VMD (Humphrey et al., 1996).

(B) Structure-based sequence alignment of the kinase domains of Rio1 and Rio2 from *A. fulgidus* (AfRio1 and AfRio2), Rio2 from *S. cerevisiae* (ScRio2), PKA, and CK1. Schematics of the secondary structure are shown in blue (Rio2) and red (CK1). Residues highlighted in red are identical in AfRio2, PKA, and CK1. Residues highlighted in green are identical in Rio2 and Rio1 sequences but not in all kinases; in blue, in Rio2 sequences only; and in yellow, in Rio1 sequences only. Phe and Tyr were colored as identical in all cases. Kinase subdomains are indicated with black Roman numerals.

Glu186, and Ser220 (Lys168 in PKA) (Figures 5A and 5C). In the ATP-bound structure, the γ-phosphate is seen in two conformations, with no contacts made with the protein (Figure 5B). In the active site of PKA, the γ-phosphate is held in place through coordination to metal ions, hydrogen bonding to Lys168 and the backbone

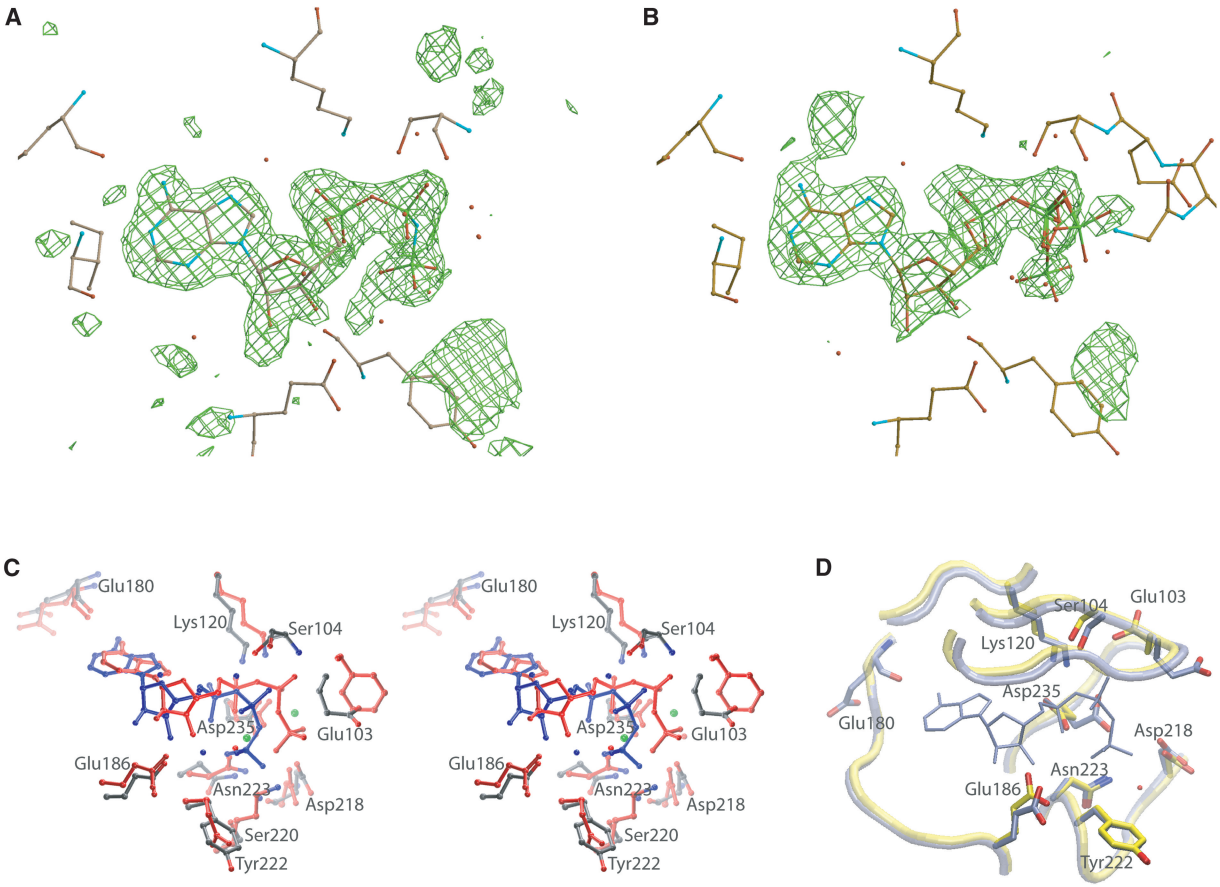

Figure 5. Nucleotide Binding by Rio2

(A) Omit map of the Rio2 active site contoured at 3 σ. $F_o - F_c$ maps were calculated after refinement of models that contained no nucleotide, using data collected from AMPPNP-soaked crystal.

(B) An analogous map calculated for an ATP-soaked crystal. The coordinates of the nucleotides resulting from the final refinements are superimposed on the maps. Figures were made using XTALVIEW (McRee, 1999) and Raster3d (Meritt and Murphy, 1994).

(C) Stereoview of the alignment of the active sites and bound nucleotides of PKA (red) and Rio2 (blue).

(D) Alignment of the active sites of the nucleotide-free Rio2 (yellow) and the AMPPNP complex (gray). (C) and (D) were prepared using VMD (Humphrey et al., 1996) and Raster3d (Meritt and Murphy, 1994).

amide of Ser53, and a water-mediated hydrogen bond to Asp166 (223 in Rio2) (Figure 5C). Therefore, there are significant differences in the positioning of the γ-phosphate and the conformation of the nucleotide in Rio2, which indicates that the crystal structure presented here may represent an inactive conformation of the enzyme.

Despite the presence of the residues expected to be important for metal binding, no metal ions were observed in either of the two nucleotide-bound structures of Rio2 presented here, or in crystals soaked with manganese ions (data not shown). Bound metals, known to be necessary for catalysis, are clearly seen in the structures of both PKA and CK1 (Knighton et al., 1991a; Xu et al., 1995). The invariant aspartic acids of the metal binding loop of Rio2, Asp235 (Asp186 in PKA) and Asn223, both seen to coordinate metal ions in PKA, are in an alternate conformation in Rio2 compared to PKA (Figure 5C). In addition, when the kinase domains of Rio2 and PKA are aligned, the loops that line the active site are significantly closer to the nucleotide in PKA than they are in Rio2 (Figure 4A). Finally, complete density was observed for the γ-phosphate in the ATP-soaked crystal despite overnight incubation, suggesting that the enzyme in the crystals was incapable of ATP hydrolysis (Figure 5B). These data further the view that this structure is of an inactive conformation of Rio2.

The Kinase and wHTH Domains of Rio2 Are Tightly Connected

The N-terminal wHTH domain and the N-lobe of the kinase domain of Rio2 are involved in extensive interdomain interactions. Ile3 from α0, Phe22 and Trp26 from the loop between α2 and α3, and Phe74, Leu77, Leu82, and Leu85 from α4 of the wHTH domain contribute to a hydrophobic core formed with Val91, Ile94, Phe107, Val119, Phe121, Ala176, and Leu178 from the β sheet and Val145 and Ile148 from the beginning of αC of the N lobe of the kinase (Figure 6). In addition, there are hydrogen bonds between Ser78, Ser81, and Asn175, as well as Arg84 and Glu173, which strengthen the interaction. Thus the wHTH domain is firmly attached to the

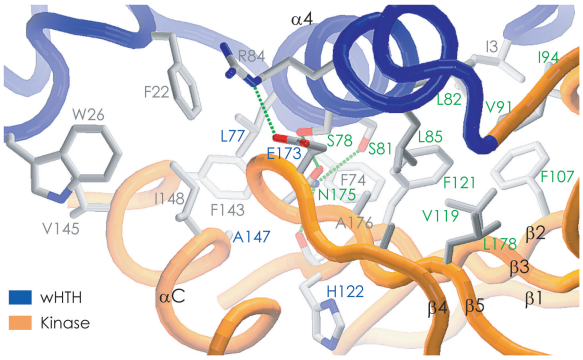

Figure 6. The Interface between the Winged Helix and the Kinase Domains

The hydrophobic packing interface of the wHTH domain (blue) and the kinase domain (orange) is shown. The residues labeled in green are chemically conserved in Rio1 and Rio2 proteins, while blue-labeled residues are chemically conserved only in Rio2 proteins. Residues labeled in gray are not conserved. Secondary structure elements are labeled in black. The figure was made using VMD (Humphrey et al., 1996) and Raster3d (Meritt and Murphy, 1994).

kinase domain, and any movement in the N lobe would result in repositioning of the wHTH domain as well.

Comparison of the Apo- and Nucleotide-Bound Structures

In most kinases, binding of a nucleotide results in a shift of the N-terminal lobe of the kinase domain relative to the C-terminal lobe (Cox et al., 1994). Rio2 is not an exception, although the observed shift is quite minimal (Figure 5D). In addition, several residues change their conformation in response to nucleotide binding. Asn223, Lys120, Glu103, and Glu186 all change positions to accommodate the nucleotide (Figure 5D). As a result of the small movement, the N-terminal domain of the protein is pulled in toward the center of the protein. These data are supported by sedimentation velocity experiments that show a decrease in the sedimentation coefficient upon nucleotide binding, which would correspond to a decrease in the Stokes' radius of the molecule (data not shown).

Conserved Residues of the RIO Proteins

The structure of Rio2 was analyzed to locate the residues that are conserved in the RIO protein family. Although many such residues are required to stabilize the tertiary structure of the protein, several surface residues are also conserved in the whole family. When mapped on the surface, the conserved residues appear to cluster around the active site on one face of the protein (Figure 7A). Tyr222 is conserved as either a Tyr or a Phe in both RIO families. Since this large, aromatic side chain is located at the mouth of the active site near the γ-phosphate of the nucleotide, it may be involved in substrate binding. Glu103 is located in the glycine-rich loop and is conserved in all RIO proteins. It moves from pointing inwards toward the center of the protein to pointing outward when nucleotide binds (Figure 5D). Lys102 is also conserved in all RIO domains but does not interact with the nucleotide and is disordered in the structure. Electrostatic analysis of the surface in this region revealed a high negative charge density surrounding a deep cleft extending away from the active site (Figure 7B). The disordered loop (residues 125–141) is located near the conserved surface and contains Gly125, Thr127, Ser128, Lys130, Asp147, and Tyr148, all of them highly conserved among Rio2 family members. Taken together, these data may indicate that the RIO domain possesses a large surface area for specific substrate binding and that the disordered loop may play a role in substrate recognition and/or in activation of the kinase.

Discussion

The structure of full-length Rio2 confirms that this family of protein kinases is distinct from the canonical serine/threonine kinases in structure as well as in sequence. No analogous combination of a wHTH domain and a kinase domain has been previously reported. In addition, the kinase domain of Rio2 is surprisingly small, making it the most compact protein kinase structurally characterized to date. This truncated version of a kinase lacks all the subdomains involved in substrate binding except for those that are part of the catalytic site. Therefore, based strictly on the structure, the specificity of this kinase for serine/threonine or tyrosine cannot be determined, although phosphoamino acid analysis performed on substrates phosphorylated with Rio1p from yeast demonstrated phosphorylation only on serine residues (Angermayr et al., 2002).

The lack of a peptide binding region in the Rio2 structure suggests that the mechanism of substrate selection may be significantly different from that seen for other serine kinases. In canonical protein serine kinases, a threonine residue (Thr201 in PKA) hydrogen bonds to the catalytic aspartate (Asp166 in PKA, 218 in Rio2) and is thought to determine the specificity of the kinase for serine/threonine rather than for tyrosine. In tyrosine kinases, a proline that is found in an equivalent position provides a nonpolar surface for the substrate tyrosine (Bossemeyer, 1995; Engh and Bossemeyer, 2002). No such residue is present in the Rio2 kinase, and the catalytic Asp218 forms no hydrogen bonds in any of the structures analyzed here. Therefore, we propose that the RIO kinase target is specific and the interaction may involve a larger surface area than seen in PKA, which has broad substrate specificity. The presence of conserved residues on the surface near the active site supports the possibility of a large surface interaction between a substrate and the kinase. The positions of some of the absolutely conserved residues, in particular Tyr222, Lys102, and Glu103, on the edge of the active site also would support a mechanism in which the kinase becomes fully activated only when substrate binding influences the conformations of these residues.

The lack of metal in the nucleotide binding site suggests that the structure presented here represents an inactive conformation of the Rio2 kinase. The disorder of the γ-phosphate and the incorrect positioning of the catalytic residues support this conclusion. Exactly what must occur to activate the Rio2 kinase is at present unknown, but since the protein is capable of autophos-

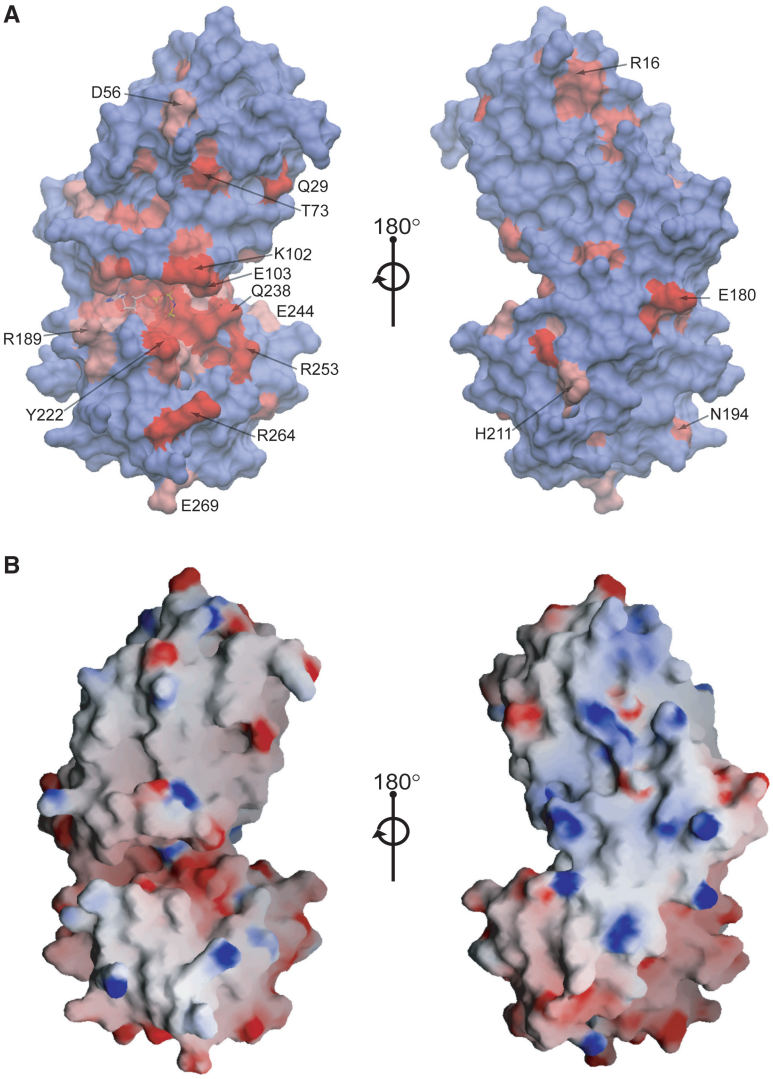

Figure 7. Conservation of Residues and Surface Properties of Rio2

(A) Surface representation of Rio2 colored according to conservation of residues. Residues identical in >95% of the sequences are red, >85% or chemically conserved are pink, and those not conserved are blue. The stick drawing shows the position of the bound nucleotide. Residues that are conserved and contain surface-exposed side chains are labeled. The figure was made using VMD (Humphrey et al., 1996).

(B) Electrostatic properties are shown for the surface of the full-length Rio2 in the same orientations as shown in (A) (−20 to +20 κBt, red to blue). The figure made using Grasp (Nicholls, 1992).

phorylation with the addition of only ATP and magnesium, we can conclude that either autophosphorylation or substrate binding may be responsible for activation of the kinase activity. Also, since the nucleotides that are present in these structures were soaked into preformed crystals, it cannot be ruled out that the constraints of crystal packing prevent conformational changes that would place the active site in an active conformation. Cocrystallization experiments attempted under the conditions that previously yielded crystals in the absence of a nucleotide were unsuccessful. A more extensive screening of crystallization conditions for the complexes is underway.

The significance of the presence of the wHTH domain is presently unknown, since neither the function of this domain nor its relationship to wHTH domains had been previously noted. Comparisons of the Rio2 wHTH domain with the DNA binding wHTH domains of transcription factors suggest that this domain is theoretically capable of binding DNA. Although most wHTH domains are involved in DNA interactions, there are examples of such domains that do not interact with DNA or are involved in RNA recognition (Dong et al., 2004). Considering its role in rRNA processing, the Rio2 wHTH domain may fall into any one of those categories.

With few exceptions, the residues conserved in the kinase domain of Rio2 that are required for correct folding of the molecule and for nucleotide binding are also conserved in Rio1 proteins. Comparison of the sequences of Rio1 and Rio2 proteins suggests that like Rio2, Rio1 lacks the residues that comprise the activation loop in typical eukaryotic protein kinases (Figure 4B). Surprisingly, the residues of the hydrophobic patch on the surface of the twisted β sheet that packs against the wHTH domain of Rio2 are conserved in Rio1 proteins (Figures 6 and 4B). This suggests that although Rio1 proteins do not contain a conserved N-terminal domain as seen in the Rio2 proteins, additional structural elements may be present N-terminal to the kinase domain that pack against this surface to complete the N-terminal lobe of the Rio1 kinase domain. The regions corresponding to the disordered loop of Rio2 (residues 125–141) are on average longer in Rio1 proteins and contain Rio1-specific conserved residues (Figure 4B). This may mean

a difference in conformation of this loop or a different role in the molecular mechanism of Rio1 proteins. An interesting difference between the two protein families is also seen in the nucleotide binding loop. In all Rio2 proteins, the consensus sequence for this loop is GV/I/EGKES, whereas in Rio1 proteins the sequence is STGKEA/S. The presence of a conserved Ser and Thr residue at the start of the loop suggests a difference in nucleotide binding and possibly in catalytic mechanism. Therefore, although there are similarities between the two families that suggest that the overall structure and possibly substrate binding may be similar, differences exists between them which could translate into differences in function, as expected from the absolute requirement for both proteins in the cell.

It must be noted that since this structure of Rio2 is from a thermophilic archaeal organism, some structural characteristics observed here may not extend to the eukaryotic proteins. The absence of the activation loop, or subdomain VIII, is expected in all RIO domains based on sequence comparison. However, the eukaryotic RIO proteins are longer on the C terminus than their counterparts in Archaea, which could mean they contain other structural elements that include the missing subdomains X and XI. Deletion of the last 80 amino acids of Rio1 in yeast is lethal and increases kinase activity in in vitro assays, suggesting a regulatory role for the C terminus (Angermayr et al., 2002). Therefore, eukaryotic RIO proteins may contain additional features evolved to control the functions of RIO proteins in eukaryotic cells.

Experimental Procedures

Protein Expression and Purification

The full-length Rio2 gene was PCR amplified from *A. fulgidus* genomic DNA (ATCC) and subcloned in pENTR-SD-Topo (Invitrogen). It was then recombined into pDEST14 using GATEWAY (Invitrogen). This construct was transformed into *E. coli* Rosetta-DE3-pLysS cells (Novagen). Expression cultures were grown at 37°C and induced with 1 mM IPTG at OD_{600} of 0.6 for 4 hr. Cells containing expressed Rio2 were harvested by centrifugation and resuspended in 50 ml of 50 mM Tris (pH 8.0), 50 mM NaCl, 0.1 X Bugbuster (Novagen), and 0.1 mg/ml DNase1 (Roche) per liter of expression culture and stirred at room temperature for 30 min. The lysate was then placed in a 75°C water bath for 15 min and immediately centrifuged to remove insoluble material. The supernatant, which contained the thermostable Rio2, was diluted 2-fold with purified water, passed through a 0.22 μM filter, and loaded onto a 5 ml HiTrap QHP column (APBiotech) equilibrated in buffer containing 25 mM Tris (pH 8.0) and 25 mM NaCl. The bound Rio2 was eluted in a gradient from 0.025 to 1 M NaCl in 20 column volumes (cv). The fractions containing Rio2 were pooled, diluted 4-fold in 25 mM Tris (pH 8.0), and loaded onto a 5 ml HiTrap Blue column (APBiotech). The protein was eluted in a gradient from 0.025 to 3 M NaCl in 20 cv. The ~98% pure protein was further purified to >99% purity by size exclusion chromatography using a column packed with Superdex 200 resin (APBiotech) equilibrated in 20 mM Tris (pH 8.0) and 200 mM NaCl. The proteins were concentrated to 5 mg/ml and stored at room temperature until crystallization screening. Se-Met Rio2 was expressed in minimal media with all amino acids supplemented except for methionine, which was replaced by Se-Met. The Se-Met protein was purified as described above with the addition of 0.2% β-mercaptoethanol in all buffers.

Kinase Assays

To assay for autophosphorylation and phosphorylation of histone H1 (Roche), myelin basic protein (MBP; Sigma-Aldrich) and dephosphorylated α-casein (Sigma-Aldrich) kinase reactions were conducted in the presence of ^{32}P-labeled ATP. The reaction buffer contained 50 mM NaCl, 50 mM Tris (pH 7.5), and 20 μCi [γ-^{32}P]ATP with or without 5 mM $MgCl_2$. All reaction mixtures for experiments shown in Figure 1 contained 2 μg of purified AfRio2 except for the second "Rio2 only" lane, which contained 10 μg Rio2, and control lanes labeled "Casein only" and "CK1 + casein", which contained no Rio2. All reactions with substrate contained 10 μg of protein substrate. The control reactions contained 8 μg of α-casein with or without 1 μg of casein kinase (CK1; Sigma-Aldrich) and 4 μCi [γ-^{32}P]ATP. The reaction mixtures were incubated at 37°C for 1 hr 30 min and run on a NuPAGE 4%–12% Bis-Tris denaturing gel (Invitrogen) for 1 hr at 120 volts. The gel was then dried and used to expose film for 5 hr for the Rio2 assays and 30 min for the CK1 control.

Crystallization

Initial crystallization conditions were obtained through sparse matrix screens (Emerald Biostructures Inc.) using the hanging drop vapor diffusion method. Diffraction quality crystals (300–500 × 50 × 50 μm in size) were obtained after 2–4 days at 20°C from drops consisting of a mixture of equal volumes of protein and well solution, containing 5%–12% polyethylene glycol (PEG) 900 and 100 mM phosphate-citrate buffer (pH 3.6–4.1) placed over 1 ml reservoirs. AMPPNP was soaked into the crystals by the addition of 2 μl of mother liquor containing 20 mM AMPPNP and 20 mM $MgCl_2$ to a 2 μl drop containing crystals in mother liquor. For the ATP-containing crystals, the same procedure was performed using 20 mM ATP, 20 mM $MgCl_2$, 100 mM Tris (pH 7.5), and 20% ethylene glycol (these crystals contained Se-Met rather than methionine). In both cases, the crystals were then incubated at 20°C for 24 hr prior to data collection.

Data Collection and Processing

Nucleotide-free, Se-Met crystals, and AMPPNP-containing crystals were flash frozen in mother liquor containing 20% ethylene glycol. ATP-containing crystals were flash frozen in 20% ethylene glycol, 100 mM Tris (pH 7.5), 20 mM ATP, 20 mM $MgCl_2$. Diffraction data for the nucleotide free, Se-Met, and AMPPNP crystals were collected at 100K at the SER-CAT beamline at the Advanced Photon Source (APS), Argonne, Illinois, with a MAR225 CCD detector. Data for the ATP-containing crystal were collected on a Rigaku RU200 rotating anode source operated at 50 kV and 100 mA, using a MAR345 image plate detector. Data from the Se-Met crystals were collected at three wavelengths (peak, inflection point, and remote). Since the space group is monoclinic, 360° of data were collected at each wavelength in 90° inverse wedges to maximize redundancy. All data were integrated and merged using HKL2000 (Otwinowski and Minor, 1997). Table 1 contains details regarding data statistics for all data sets.

Structure Determination and Refinement

The autoSharp program suite was used for the Se-Met data set to obtain the phases using the MAD method, apply solvent flattening and density modification to the initial electron density map, and perform automatic model building with wARP (Perrakis et al., 1999). The almost complete model (which, however, was not fully refined) was used to phase the native data set containing no nucleotide, and the latter structure was finalized by rebuilding in XTALVIEW (McRee, 1999) and refinement with REFMAC5 (Murshudov et al., 1997). The completely refined model was used to phase the AMPPNP- and ATP-containing data sets, and those were submitted to several rounds of building in XTALVIEW and refinement using REFMAC5. R_{free} was monitored by using 5% of the reflections as a test set in each case. Refinement statistics are provided in Table 1.

Acknowledgments

We are grateful to Sook M. Lee and Peter F. Johnson, NCI-Frederick, for assistance with kinase assays, and Andrew G. Stephen and Robert J. Fisher, Protein Chemistry Laboratory, SAIC Frederick, Inc., for DNA binding studies. Diffraction data were collected at the Southeast Regional Collaborative Access Team (SER-CAT) beamline 22-ID, located at the Advanced Photon Source, Argonne National Laboratory. Use of the Advanced Photon Source was sup-

ported by the U. S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. W-31-109-Eng-38.

Received: May 15, 2004
Revised: June 11, 2004
Accepted: June 12, 2004
Published: September 7, 2007

References

Alekshun, M.N., Levy, S.B., Mealy, T.R., Seaton, B.A., and Head, J.F. (2001). The crystal structure of MarR, a regulator of multiple antibiotic resistance, at 2.3 Å resolution. Nat. Struct. Biol. *8*, 710–714.

Angermayr, M., and Bandlow, W. (2002). RIO1, an extraordinary novel protein kinase. FEBS Lett. *524*, 31–36.

Angermayr, M., Roidl, A., and Bandlow, W. (2002). Yeast Rio1p is the founding member of a novel subfamily of protein serine kinases involved in the control of cell cycle progression. Mol. Microbiol. *44*, 309–324.

Bossemeyer, D. (1995). Protein kinases–structure and function. FEBS Lett. *369*, 57–61.

Brown, N.R., Noble, M.E., Lawrie, A.M., Morris, M.C., Tunnah, P., Divita, G., Johnson, L.N., and Endicott, J.A. (1999). Effects of phosphorylation of threonine 160 on cyclin-dependent kinase 2 structure and activity. J. Biol. Chem. *274*, 8746–8756.

Carson, M. (1991). RIBBONS 4.0. J. Appl. Crystallogr. *24*, 958–961.

Chen, P., Luo, C., Deng, Y., Ryan, K., Register, J., Margosiak, S., Tempczyk-Russell, A., Nguyen, B., Myers, P., Lundgren, K., et al. (2000). The 1.7 A crystal structure of human cell cycle checkpoint kinase Chk1: implications for Chk1 regulation. Cell *100*, 681–692.

Clark, K.L., Halay, E.D., Lai, E., and Burley, S.K. (1993). Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5. Nature *364*, 412–420.

Cox, S., Radzio-Andzelm, E., and Taylor, S.S. (1994). Domain movements in protein kinases. Curr. Opin. Struct. Biol. *4*, 893–901.

Dong, G., Chakshusmathi, G., Wolin, S.L., and Reinisch, K.M. (2004). Structure of the La motif: a winged helix domain mediates RNA binding via a conserved aromatic patch. EMBO J. *23*, 1000–1007.

Engh, R.A., and Bossemeyer, D. (2002). Structural aspects of protein kinase control-role of conformational flexibility. Pharmacol. Ther. *93*, 99–111.

Gajiwala, K.S., and Burley, S.K. (2000). Winged helix proteins. Curr. Opin. Struct. Biol. *10*, 110–116.

Gajiwala, K.S., Chen, H., Cornille, F., Roques, B.P., Reith, W., Mach, B., and Burley, S.K. (2000). Structure of the winged-helix protein hRFX1 reveals a new mode of DNA binding. Nature *403*, 916–921.

Gavin, A.C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J.M., Michon, A.M., Cruciat, C.M., et al. (2002). Functional organization of the yeast proteome by systematic analysis of protein complexes. Nature *415*, 141–147.

Geerlings, T.H., Faber, A.W., Bister, M.D., Vos, J.C., and Raue, H.A. (2003). Rio2p, an evolutionarily conserved, low abundant protein kinase essential for processing of 20 S Pre-rRNA in Saccharomyces cerevisiae. J. Biol. Chem. *278*, 22537–22545.

Giaever, G., Chu, A.M., Ni, L., Connelly, C., Riles, L., Veronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., Andre, B., et al. (2002). Functional profiling of the *Saccharomyces cerevisiae* genome. Nature *418*, 387–391.

Hanks, S.K., and Hunter, T. (1995). Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. FASEB J. *9*, 576–596.

Hanks, S.K., Quinn, A.M., and Hunter, T. (1988). The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. Science *241*, 42–52.

Ho, Y., Gruhler, A., Heilbut, A., Bader, G.D., Moore, L., Adams, S.L., Millar, A., Taylor, P., Bennett, K., Boutilier, K., et al. (2002). Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. Nature *415*, 180–183.

Holm, L., and Sander, C. (1993). Protein structure comparison by alignment of distance matrices. J. Mol. Biol. *233*, 123–138.

Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD: visual molecular dynamics. J. Mol. Graph. *14*, 33–38.

Knighton, D.R., Zheng, J.H., Ten Eyck, L.F., Ashford, V.A., Xuong, N.H., Taylor, S.S., and Sowadski, J.M. (1991a). Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. Science *253*, 407–414.

Knighton, D.R., Zheng, J.H., Ten Eyck, L.F., Xuong, N.H., Taylor, S.S., and Sowadski, J.M. (1991b). Structure of a peptide inhibitor bound to the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. Science *253*, 414–420.

Leonard, C.J., Aravind, L., and Koonin, E.V. (1998). Novel families of putative protein kinases in bacteria and archaea: evolution of the "eukaryotic" protein kinase superfamily. Genome Res. *8*, 1038–1047.

Manning, G., Whyte, D.B., Martinez, R., Hunter, T., and Sudarsanam, S. (2002). The protein kinase complement of the human genome. Science *298*, 1912–1934.

McRee, D.E. (1999). XtalView/Xfit: a versatile program for manipulating atomic coordinates and electron density. J. Struct. Biol. *125*, 156–165.

Meritt, E.A., and Murphy, M.E.P. (1994). Raster3D Version 2.0. A program for photorealistic molecular graphics. Acta Crystallogr. D *50*, 869–873.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D *53*, 240–255.

Nicholls, A. (1992). GRASP: Graphical Representation and Analysis of Surface Properties (New York: Columbia University).

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. *276*, 307–326.

Perrakis, A., Morris, R., and Lamzin, V.S. (1999). Automated protein model building combined with iterative structure refinement. Nat. Struct. Biol. *6*, 458–463.

Ruggero, D., and Pandolfi, P.P. (2003). Does the ribosome translate cancer? Nat. Rev. Cancer *3*, 179–192.

Schafer, T., Strauss, D., Petfalski, E., Tollervey, D., and Hurt, E. (2003). The path from nucleolar 90S to cytoplasmic 40S pre-ribosomes. EMBO J. *22*, 1370–1380.

Vanrobays, E., Gleizes, P.E., Bousquet-Antonelli, C., Noaillac-Depeyre, J., Caizergues-Ferrer, M., and Gelugne, J.P. (2001). Processing of 20S pre-rRNA to 18S ribosomal RNA in yeast requires Rrp10p, an essential non-ribosomal cytoplasmic protein. EMBO J. *20*, 4204–4213.

Vanrobays, E., Gelugne, J.P., Gleizes, P.E., and Caizergues-Ferrer, M. (2003). Late cytoplasmic maturation of the small ribosomal subunit requires RIO proteins in *Saccharomyces cerevisiae*. Mol. Cell. Biol. *23*, 2083–2095.

Wolberger, C., and Campbell, R. (2000). New perch for the winged helix. Nat. Struct. Biol. *7*, 261–262.

Xu, R.M., Carmel, G., Sweet, R.M., Kuret, J., and Cheng, X. (1995). Crystal structure of casein kinase-1, a phosphate-directed protein kinase. EMBO J. *14*, 1015–1023.

Zheng, J., Knighton, D.R., Ten Eyck, L.F., Karlsson, R., Xuong, N., Taylor, S.S., and Sowadski, J.M. (1993). Crystal structure of the catalytic subunit of cAMP-dependent protein kinase complexed with MgATP and peptide inhibitor. Biochemistry *32*, 2154–2161.

Accession Numbers

Final coordinates and structure factors have been submitted to the Protein Data Bank (accession codes 1TQI, 1TQM, and 1TQP for the Apo, AMPPNP-, and ATP-bound enzyme, respectively).